CLAIMS

What is claimed is:

1. A method of training acoustic models of a segmentation-based automatic speech recognition system, comprising the steps of:

receiving correct alignment data that represents a correct segment alignment of an utterance that was received by the speech recognition system;

receiving wrong alignment data that represents an alignment of the utterance that is known to be incorrect based on information received from the speech recognition system and describing the utterance;

identifying a first phoneme in the wrong alignment data that corresponds to a second phoneme in the correct alignment data;

modifying a first acoustic model of the first phoneme by moving at least one mean value thereof further from the feature values used to score the first phoneme.

2. A method as recited in Claim 1, further comprising the steps of:

receiving correct alignment data that represents an alignment of the utterance that is known to be correct based on information received from the speech recognition system and describing the utterance;

identifying a second phoneme in the correct alignment data that corresponds to the first phoneme in the wrong alignment data;

modifying a second acoustic model of the second phoneme by moving at least one mean value thereof closer to the feature values used to score the second phoneme.

3. A method as recited in Claim 1, wherein receiving correct alignment data comprises the step of receiving correct alignment data that represents a segment alignment of a highest scoring hypothesized alignment selected from among n-best hypotheses of an utterance that was received by the speech recognition system.

4. A method as recited in Claim 1, wherein receiving wrong alignment data comprises the steps of receiving wrong alignment data that represents an alignment of the utterance that is known to be incorrect based on user confirmation information received from the speech recognition system in response to prompting a speaker to confirm the utterance.

5. A method as recited in Claim 1, wherein receiving correct alignment data comprises the steps of receiving correct alignment data that represents an alignment of the utterance that is known to be correct based on user confirmation information received from the speech recognition system in response to prompting a speaker to confirm the utterance.

6. A method as recited in Claim 1, further comprising the step of iteratively repeating the identifying and modifying steps for all phonemes in the wrong alignment data that correspond to one or more phonemes in the correct alignment data.

7. A method as recited in Claim 2, further comprising the step of iteratively repeating the identifying and modifying steps for all phonemes in the correct alignment data that correspond to one or more phonemes in the wrong alignment data.

8. A method as recited in Claim 1, wherein the step of moving at least one mean value further from a corresponding mean value of a second acoustic model of the second phoneme comprises subtracting the mean value of the third acoustic model from the mean value of the second acoustic model.

9. A method as recited in Claim 1, wherein the step of moving at least one mean value further from a corresponding mean value of a second acoustic model of the second phoneme comprises reducing the mean value of the third acoustic model by approximately two percent (2%).

10. A method as recited in Claim 1, wherein modifying a first acoustic model further comprises the steps of modifying all acoustic models associated with the first phoneme by moving all mean values thereof further from corresponding mean values of all second acoustic models associated with the second phoneme.

11. A method as recited in Claim 2, wherein modifying a third acoustic model further comprises the steps of modifying all acoustic models associated with the third phoneme by moving all mean values thereof closer to corresponding mean values of all acoustic models associated with the second phoneme

12. A method of improving performance of a segmentation-based automatic speech recognition system (ASR) by training its acoustic models using information obtained from a particular application in which the ASR is used, comprising the steps of:

receiving a correct segment alignment of an utterance that was received by the ASR;

receiving an alignment of the utterance that is known to be incorrect based on information received from the speech recognition system in the context of the particular application;

identifying a first phoneme in the known incorrect alignment that corresponds to a second phoneme in the correct segment alignment;

modifying a first acoustic model of the first phoneme by moving at least one mean value thereof further from a corresponding mean value of a second acoustic model of the second phoneme.

13. A method as recited in Claim 12, further comprising the steps of:

receiving an alignment of the utterance that is known to be correct based on information received from the speech recognition system in the context of the particular application;

identifying a third phoneme in the known correct alignment that corresponds to the second phoneme in the correct alignment;

modifying a third acoustic model of the third phoneme by moving at least one mean value thereof closer to the corresponding mean value of the second acoustic model of the second phoneme.

14. A computer-readable medium carrying one or more sequences of instructions for training acoustic models of a segmentation-based automatic speech recognition system, wherein execution of the one or more sequences of instructions by one or more processors causes the one or more processors to perform the steps of:

receiving correct alignment data that represents a correct segment alignment of an utterance that was received by the speech recognition system;

receiving wrong alignment data that represents an alignment of the utterance that is known to be incorrect based on information received from the speech recognition system and describing the utterance;

identifying a first phoneme in the wrong alignment data that corresponds to a second phoneme in the correct alignment data;

modifying a first acoustic model of the first phoneme by moving at least one mean value thereof further from a corresponding mean value of a second acoustic model of the second phoneme.

15. A computer-readable medium as recited in Claim 14, wherein the instructions further comprise instructions for carrying out the steps of:

receiving an alignment of the utterance that is known to be correct based on information received from the speech recognition system in the context of the particular application;

identifying a third phoneme in the known correct alignment that corresponds to the second phoneme in the correct alignment;

modifying a third acoustic model of the third phoneme by moving at least one mean value thereof closer to the corresponding mean value of the second acoustic model of the second phoneme.

16. A segmentation-based automatic speech recognition system that provides improved performance by training its acoustic models according to information about an application with which the system is used, comprising:

a recognizer that includes one or more processors;

non-volatile storage coupled to the recognizer and comprising a plurality of segmentation alignment data and a plurality of acoustic models;

a computer-readable medium coupled to the recognizer and carrying one or more sequences of instructions for the training acoustic models, wherein execution of the one or more sequences of instructions by the one or more processors causes the one or more processors to perform the steps of:

receiving correct alignment data that represents a correct segment alignment of an utterance that was received by the speech recognition system;

receiving wrong alignment data that represents an alignment of the utterance that is known to be incorrect based on information received from the speech recognition system and describing the utterance;

identifying a first phoneme in the wrong alignment data that corresponds to a second phoneme in the correct alignment data;

modifying a first acoustic model of the first phoneme by moving at least one mean value thereof further from a corresponding mean value of a second acoustic model of the second phoneme.

17. A speech recognition system as recited in Claim 16, wherein the instructions further comprise instructions for carrying out the steps of:

receiving an alignment of the utterance that is known to be correct based on information received from the speech recognition system in the context of the particular application;

identifying a third phoneme in the known correct alignment that corresponds to the second phoneme in the correct alignment;

modifying a third acoustic model of the third phoneme by moving at least one mean value thereof closer to the corresponding mean value of the second acoustic model of the second phoneme.